Multiple meaning tokens may be associated with each of one or more polysemous vocabulary words contained in the speech recognition dictionary.

The automatic speech recognition system may include a language analyzer that is configured to extract meaning from the sequence of meaning tokens provided by the speech recognizer based upon a set of task-specific semantic rules. The language analyzer may be a deterministic rule-based language analyzer. The automatic speech recognition system also may include an application command translator that is configured to select an action from a set of application-specific actions based upon the meaning extracted by the language analyzer, and to issue one or more commands to carry out the selected action.

The speech recognition dictionary preferably is a data structure that is stored in a computer-readable physical medium.

The invention also features an automatic speech recognition method in accordance with which spoken input is converted into a sequence of meaning tokens that are contained in a speech recognition dictionary and correspond to a sequence of vocabulary words that are most likely to have been spoken by a user. The speech recognition dictionary comprises a plurality of meaning tokens each associated with one or more pronunciations of one or more vocabulary words and signifying a single meaning.

The invention also features a computer program for automatically recognizing speech. The computer program resides on a computer-readable medium and comprises computer-readable instructions for causing a computer to convert spoken input into a sequence of meaning tokens that are contained in a speech recognition dictionary and corresponding to a sequence of vocabulary words that are most likely to have been spoken by a user. The speech recognition dictionary resides on the computer-readable medium and comprises a plurality of meaning tokens each associated with one or more pronunciations of one or more vocabulary words and signifying a single meaning.

Among the advantages of the invention are the following.

An automatic speech recognizer may function normally with respect to the inventive meaning token dictionary. In particular, the invention readily may be

incorporated into existing automatic speech recognition systems in which the dictionaries may be manipulated or replaced. The output from the speech recognizer, however, is in the form of meaning tokens rather than simple transcriptions of ordinary spoken words. The inventive meaning token dictionary may simplify the work that must be performed by the language analyzer and any downstream application program that uses an automatic speech recognition system for spoken language input. In addition, the output from the speech recognition system may be substantially free of ambiguities, allowing the downstream application program to use deterministic interpretation algorithms (which are simpler and more efficient than non-deterministic algorithms).

Because the meaning tokens may have longer pronunciations than ordinary words, more data may be used to distinguish words in the speech recognition dictionary. In this way, the invention enables the speech recognition system to recognize each word with greater precision, resulting in higher overall recognition accuracy. In addition, since there may be fewer word candidates with pronunciations that are similar enough to the sound input, there may be fewer hypotheses to test against the grammar, increasing the efficiency of the automatic speech recognition process.

The grammars that may be used with the invention also may be much simpler than grammars that, for example, must account for all combinations of words. For example, idiomatic and context dependent word combinations, such as "go for it" or "look it up," may be handled in the inventive meaning token dictionary rather than in the grammar.

Other features and advantages of the invention will become apparent from the following description, including the drawings and the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 is a block diagram of an automatic speech recognition system that includes a speech recognizer, a language analyzer, and an application command translator.

FIG. 2 is a block diagram of a computer on which the automatic speech recognition system of FIG. 1 may be implemented.

FIG. 3 is a block diagram of the speech recognizer of FIG. 1.

FIG. 4 is a block diagram of a meaning token dictionary associating a pronunciation with a meaning token having dictionary spelling signifying a single meaning.

FIG. 5 is a block diagram of a meaning token dictionary associating multiple pronunciations with a single meaning token.

FIG. 6 is a block diagram of a meaning token dictionary associating a single pronunciation with one of multiple possible meaning tokens.

DETAILED DESCRIPTION

In the following description, like reference numbers are used to identify like elements. Furthermore, the drawings are intended to illustrate major features of exemplary embodiments in a diagrammatic manner. The drawings are not intended to depict every feature of actual embodiments nor relative dimensions of the depicted elements, and are not drawn to scale.

Referring to FIG. 1, in one embodiment, an automatic speech recognition system 10 includes a speech recognizer 12, a language analyzer 14, and an application command translator 16. In operation, speech recognizer 12 converts an incoming stream of speech sounds 18 (or utterances) into a likely sequence of meaning tokens 20. In particular, speech recognizer 12 receives digitized speech samples from an acoustic input device (e.g., a microphone; not shown), and converts the digitized speech samples into sequences of recognized meaning tokens based upon finite state machine templates. The finite state machine templates are defined by a set of vocabulary meaning token patterns, which are stored in a meaning token dictionary. In some embodiments, the finite state machine templates also are defined by a set of grammar rules. Language analyzer 14 may be implemented as a conventional deterministic rule-based language analyzer. Language analyzer 14 attempts to make sense of the meaning token sequences 20 that are output by speech recognizer 12. That is, language analyzer 14 attempts to extract from each meaning